Screening for Voice Disorders in Older Adults (*Rastreamento de Alterações Vocais em Idosos—RAVI*)—Part I: Validity Evidence Based on Test Content and Response Processes

^{*},[†]Leandro de Araújo Pernambuco, [‡],[§]||Albert Espelt, ^{*},[†]Hipólito Virgílio Magalhães Júnior, [†]Renata Veiga Andersen Cavalcanti, and ^{*}Kenio Costa de Lima, ^{*}†Rio Grande do Norte, Brazil, and [‡]Barcelona, [§]Bellaterra, ||Madrid, Spain

Summary: Purpose. To identify the validity evidence based on the content and response processes of the *Rastreamento de Alterações Vocais em Idosos* (RAVI; “Screening for Voice Disorders in Older Adults”), an epidemiologic screening for voice disorders in older adults.

Study Design. This is a prospective, nonrandomized, cross-sectional, validation study.

Methods. Criteria for defining the domains and elaborating the questions were established to confirm the validity evidence based on the content. A multidisciplinary committee of 19 experts evaluated the questions, and the relevance and representation of the domains were analyzed using the Content Validity Index for Items (CVI-I) and the Content Validity Index (CVI), respectively. For validity evidence based on the response processes, 40 individuals of both sexes, aged ≥ 60 years, were stratified by demographic and socioeconomic condition. They responded to the RAVI, made suggestions, and their verbal and nonverbal reactions were observed.

Results. The first version of the RAVI consisted of 20 questions related to sensations and perceptions associated with the voice. Although the CVI value of 0.80 was satisfactory, the CVI-I and the suggestions of the expert committee indicated that the scale needed to be reformulated. Consultation with older adults indicated a need for further adjustment. The preliminary version of the RAVI consisted of 16 questions.

Conclusions. The two aspects of validity evidence described in the present study were essential for adapting the questions to better fit the construct of the questionnaire. Other aspects of validity evidence and reliability analysis will be described in part II of this study.

Key Words: Voice–Voice disorders–Dysphonia–Aged–Aging–Health of the elderly–Epidemiology–Validation studies.

INTRODUCTION

Aging can bring a decline in various aspects involved in the ability to communicate, such as breathing, articulation, language, hearing, and speech.¹ With the exponential increase in the number of older adults around the world, it is to be expected that changes in the quality of communication of older adults are becoming more frequent, including those specifically related to the voice.^{1,2}

The voice is an essential element of human communication,³ being the primary auditory component of speech,⁴ and allows ideas, emotions, and feelings to be shared and expressed⁵ in personal, social, and professional relationships. Just like any other physiological mechanism of the human body, the voice

is susceptible to the effects of aging, whether natural or resulting from the greater exposure to disease at this age.^{6–8}

Older adults’ voices are characterized by hoarseness, breathiness, pitch changes, difficulty in controlling loudness, restricted vocal modulation, diminished vocal resistance, instability, effort, and vocal fatigue.^{6–9} The severity and intensity of these characteristics may vary according to the person and are related to the reduced mobility and biomechanical efficiency of the entire laryngeal system that result from age-related natural morphologic and physiological changes. Furthermore, neurologic and chronic diseases affecting the voice occur more frequently in old age and may result in a worsening of the natural vocal characteristics or the appearance of specific vocal symptoms that, in some cases, may even be early signs of functional impairment.^{7,8,10,11}

Voice disorders can negatively affect the quality of life, degree of autonomy, social integration, and physical, emotional, and social skills of older adults. This, in turn, increases the risk of disability, dependency, social isolation, poor productivity, disease, and a decline in general health.^{4,5,9–12} Some studies^{13,14} have revealed that even if older people recognize changes in their voice due to aging, they are generally unaware of treatment options and accept these vocal changes as a natural part of old age. Researchers have called attention to the need for early identification and intervention of voice disorders in this population.^{11,13,14}

Accepted for publication April 10, 2015.

From the ^{*}Programa de Pós-graduação em Saúde Coletiva (PPGSCol-UFRN), Universidade Federal do Rio Grande do Norte (UFRN), Rio Grande do Norte, Brazil; [†]Department of Speech, Language and Hearing Sciences, Universidade Federal do Rio Grande do Norte (UFRN), Rio Grande do Norte, Brazil; [‡]Agència de Salut Pública de Barcelona (ASPB), Servei d’Atenció i Prevenció a les Drogodependències, Barcelona, Spain; [§]Departament de Psicobiologia i Metodologia de les Ciències de la Salut, Facultat de Psicologia, Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain; and the ||CIBER de Epidemiologia y Salud Pública, Instituto de Salud Carlos III, Madrid, Spain.

Address correspondence and reprint requests to Leandro de Araújo Pernambuco, Av. General Gustavo Cordeiro de Farias, S/N, Petrópolis, Natal, RN 59012-570, Brazil. E-mail: leandroape@globo.com

Journal of Voice, Vol. 30, No. 2, pp. 246.e9-246.e17
0892-1997/\$36.00

© 2016 The Voice Foundation

<http://dx.doi.org/10.1016/j.jvoice.2015.04.008>

Population-based studies estimate that the prevalence of vocal disorders among older adults varies from 4.8% to 29.1%.^{11,15} However, the heterogeneity of methods used in such studies compromises the accuracy of their estimates, in part due to the lack of a short, inexpensive, valid, reliable, and easily administered tool for use in health inquiries to determine the population prevalence of voice disorders among older adults.^{11,13–15} There are a number of instruments available in the literature designed to evaluate vocal handicaps or the impact of voice disorders on quality of life or to monitor the results of therapeutic intervention in adults. Such instruments include the Voice-related Quality of Life,¹⁶ Voice Handicap Index,¹⁷ Voice Symptom Scale,¹⁸ Voice Activity and Participation Profile,¹⁹ and Glottal Function Index.²⁰ However, none of these tools can be used for epidemiologic diagnostic purposes or were specifically designed for older adults. This population would have specific domains and experiences regarding voice that would differ from those of younger adults. The Geriatric Index of Communicative Ability,¹ currently undergoing validation, is an instrument for investigating aspects of communicative impairment in older adults; however, despite including questions related to the voice, it is not intended to specifically identify voice disorders. As such, we propose a new questionnaire to fill this gap in voice-disorder scales, the *Rastreamento de Alterações Vocais em Idosos* (RAVI; “Screening for Voice Disorders in Older Adults”), a self-reported outcome questionnaire for epidemiologic detection of voice disorders in older adults.

For the RAVI to be both epidemiologically relevant and scientifically robust, it must be formally developed and psychometrically tested. Thus, it is necessary to investigate its validity (ie, whether it measures what it intends to measure) and reliability (ie, whether it produces consistent and reproducible data).²¹ The most current edition of the “Standards for Educational and Psychological Testing,”²¹ the document outlining the most robust, traditional, and frequently used guidelines relating to the development and interpretation of tests, specifies, in addition to reliability, five sources of validity evidence, those based on content, response processes, internal structure, relation to other variables, and test consequences.

Despite being strongly recommended, the validity evidence based on the test content and response processes is rarely obtained and presented in detail in literature.²² According to the “Standards,”²¹ validity evidence based on content “can be obtained from an analysis of the relationship between a test’s content and the construct it is intended to measure.” This step is meant to show that the content of a test or questionnaire is suitable for evaluating a particular phenomenon, such as the presence of voice disorders. Also according to the “Standards,”²¹ “test content refers to the themes, wording, and format of the items, tasks, or questions on a test, as well as the guidelines for procedures regarding administration and scoring.” Finally, response process refers to “evidence concerning the fit between the construct and the detailed nature of performance or response actually engaged in by examinees.”²¹ The response process is

the result of observations of the administration of the test or questionnaire (eg, time to respond) and examining participants’ behavior during this time (eg, fatigue while responding).

Regarding voice research, it has been reported that there were some problems in the validation process of existing self-assessment instruments, due in part to failings in the development phase of these instruments.²³ We agree that the initial stages of the development of any instrument should be reported; as such, we chose to present the entire process of obtaining validity evidence and reliability in two articles. Part I, presented herein, describes how the validity evidence based on the questionnaire content and response processes was obtained.

METHODS

Validity evidence based on test content

The procedure to obtain validity evidence based on test content involved two key steps: development of questions and assessment of questions by a committee of experts.²⁴ The results obtained in these two steps were analyzed according to the elements that evaluate the quality of an instrument’s content: domain definition, domain representation, and domain relevance.^{25,26} Domain definition “provides the details regarding what the test measures and so it transforms the theoretical construct to a more concrete content domain.”²⁶ In the case of the RAVI, it was necessary to first clarify what the questionnaire should measure and, from that, set the domain. Domain representation “refers to the degree to which a test adequately represents and measures the domain as defined in the test specifications.”²⁶ Thus, it was necessary to observe whether the RAVI really represented the previously defined domain. Finally, domain relevance “addresses the extent to which each item on a test is relevant to the targeted domain,”²⁶ that is, each question is analyzed individually to examine whether it is relevant to the previously defined domain.

Development of questions

The conceptual definition of the instrument’s construct and the initial preparation of questions were derived from available scientific literature (theoretical basis) and the clinical experience and observations (empirical basis) of four researchers with expertise in the construct (researchers included the first, third, fourth, and fifth authors of this article). These four researchers met initially to conceptually define the construct “vocal disorders.” Following definition, researchers brainstormed on what content they considered relevant for the questionnaire. Subsequently, the contents were filtered and transformed into questions.

Questions relating to risk factors, redundant questions, and questions that did not fit the age range of the target population were avoided by consensus. Regarding syntactic and semantic definition, the researchers considered the educational level of the population, whether the questions were too long, and the presence of ambiguous words, vague terms, jargon, and questions that covered two or more aspects of the construct at the same time. This initial version consisted of 20 questions

TABLE 1.
Distribution Criteria of Four Social Groups Used to Evaluate the Instrument in Terms of the Validity Evidence Based on Response Processes

Criteria	Group A		Group B		Group C			Group D
Education	HEC	HSC	HSI	HEC	PSI to HSI	HSC	HEI	No formal education to HSC
Private health care system	Y or N	Y	Y or N	Y or N	Y	Y	Y or N	N
Income according to neighborhood of residence*	T3	T3 or T2	T3 or T2	T2 or T1	T3, T2, or T1	T1	T1	T3, T2, or T1
Male	4		4			4		4
Female	6		6			6		6
Age								
Mean (standard deviation)	70.6 (±6.09)		68.2 (±4.70)			74.8 (±5.51)		73.2 (±5.55)
Minimum–maximum	62–81		63–78			65–83		66–83

Abbreviations: HEC, higher education complete; HSC, high school complete; HSI, high school incomplete; PSI, primary school incomplete; HEI, higher education incomplete; Y, yes; N, no; T3, upper tercile; T2, medium tercile; T1, lower tercile.

* This criterion was categorized using nominal mean monthly earnings (in “reais,” the currency of Brazil) of individuals aged ≥60 years who resided in the neighborhoods of Natal and Parnamirim (Rio Grande do Norte, Brazil) according to information from the 2010 Census, which was obtained from the “Instituto Brasileiro de Geografia e Estatística” (IBGE) site. The absolute earnings values were categorized into terciles and formed three strata (T1, T2, and T3).

associated with sensations and perceptions related to the voice. For this study, “sensation” questions referred to any self-report of a physical reaction, whereas “perception” questions involved perceptive self-assessment of voice elements that older adults could capture during production of their own voices.

When an older adult attested to the presence of a sensation or perception, he or she had to rate that sensation or perception on another four scales: time of onset, frequency (defined as how often it occurs—less frequent, frequent, or very frequent; scored by the researchers from one to three, respectively), intensity (defined as how strong it is—weak, moderate, or intense; scored by the researchers from one to three, respectively), and severity (defined as the intersection between frequency and intensity—mild, moderate, or severe; scored by the researchers from one to three, respectively). The results for severity were not directly solicited from older adults but were the result of multiplying the scores of frequency and intensity.

Evaluation of questions by a committee of experts

A multidisciplinary committee of experts evaluated the questions. To be considered an expert, a professional had to have ≥5 years of experience and recognized expertise in treatment of voice disorders, voice research, or gerontology. This committee was different from the panel that generated the questions. Each expert evaluated whether each question was well developed, relevant to the purpose of the questionnaire, and understandable to older adults. The experts had the opportunity to recommend new questions or suggest modifications to existing questions and provided their comments individually without consulting each other. This step was based on the conceptual

premise of the Delphi method, which is used to capture and structure the opinions, judgments, and expertise of a panel of people who have expertise in a particular subject.^{27,28} In the present study, the conceptual framework of the Delphi method and a single round of consultation with each expert were used.^{27,29}

The committee of experts comprised 19 professionals, including 13 speech-language therapists (including gerontologists, specialists in gerontology, voice experts, and professionals with recognized experience in the field of aging), two geriatricians, two psychologists, one social worker, and one physiotherapist. In the instrument sent to experts via e-mail, the options “adequate” and “inadequate” were available for each question. In addition, there was a space for justifications, suggestions, adjustments, or the indication of new questions.

To analyze the level of agreement between experts on each individual question (domain relevance) and the instrument as a whole (domain representation), the Content Validity Index by Item (CVI-I; the number of experts who judged each question as “adequate” divided by the total number of judgments) and the Content Validity Index (CVI; mean of the CVI-I) were used,³⁰ respectively. Questions with a result below the reference value of 0.78³⁰ were excluded or modified. Finally, the authors of the questionnaire met to discuss the relevant reformulations and the possible incorporation of the opinions and suggestions of the experts.

To strengthen the validity evidence based on the content, the suggestions of the target population were obtained. Forty older adults were encouraged to suggest possible modifications or propose relevant aspects that should be considered but had not been included in the questionnaire. This group of older individuals was the same as that that provided data

TABLE 2.
Results Obtained From the Committee of Experts During the Assessment of the Validity Evidence Based on the Content

Question	Content Validity by Item Index (CVI-I)	Decision of Authors	Order of Question in Instrument After Adjustments
1. Você sente que tem rouquidão? (Do you feel that you have hoarseness?)	0.63	Change to “Sua voz é rouca?” (“Is your voice hoarse?”)—suggestion of experts	3°
2. Você se sente incomodado com a sua voz? (Do you feel uncomfortable with your voice?)	0.84	Change to “Sua voz lhe incomoda?” (“Does your voice bother you?”)—adjustment by authors	1°
3. Você sente que as pessoas se incomodam com a sua voz? (Do you feel that people are bothered by your voice?)	0.94	Change to “As pessoas se incomodam com sua voz?” (“Does your voice bother other people?”)—adjustment by authors	2°
4. Você sente que faz esforço para falar? (Do you feel that you have to make an effort to speak?)	0.84	Change to “Faz esforço para falar?” (“Do you make an effort when speaking?”)—adjustment by authors	12°
5. Você se cansa para falar? (Does talking make you tired?)	0.84	Change to “Sente cansaço ao falar?” (“Do you feel tired when speaking?”)—adjustment by authors	13°
6. Você sente que fala fraco/baixo ou forte/alto? (Do you feel that you speak weakly/quietly or strongly/loudly?)	0.73	Separate into two questions: “Sua voz é fraca ou baixa?” (“Is your voice weak or quiet?”) and “Sua voz é forte ou alta?” (“Is your voice strong or loud?”)	4° e 5°
7. Você sente dificuldade para falar fraco/baixo ou forte/alto? (Do you experience difficulty when you speak weakly/quietly or strongly/loudly?)	0.78	Delete (two extremes in a single question can potentially create confusion with question 06)	—
8. Você sente que a sua voz tem entonação? (Do you feel that your voice has intonation?)	0.21	Delete (low relevance for the construct and not widely applicable)	—
9. Você sente que a sua voz treme? (Do you feel that your voice is shaking?)	0.94	Change to “Percebe tremor na voz?” (“Do you notice your voice shaking?”)—adjustment by authors	6°
10. Você sente que sua voz está mais fina (aguda) ou mais grossa (grave)? (Do you feel that your voice is high pitched or low pitched?)	0.78	Separate into two questions: “Sua voz tem ficado mais fina (aguda)?” (“Has your voice become high pitched?”) and “Sua voz tem ficado mais grossa (grave)?” (“Has your voice become low pitched?”)	7° e 8°
11. Você sente dificuldade para usar a voz ao cantar? (Do you have difficulty using your voice to sing?)	0.89	Delete (not a common activity for all those in the target population)	—
12. Você sente sua garganta seca? (Do you feel that your throat is dry?)	0.89	Change to “Sente sua garganta seca ao falar?” (“Do you feel that your throat is dry when you speak?”)—adjustment by authors	14°
13. Você sente dor na garganta enquanto fala? (Do you feel pain in the throat while talking?)	0.89	Change to “Sente dor na garganta ao falar?” (“Do you feel pain in the throat when you speak?”)—adjustment by authors	15°
14. Você sente queimação na garganta? (Do you feel burning in the throat?)	0.84	Change to “Sente queimação na garganta ao falar?” (“Do you feel a burning sensation in the throat when you speak?”)—adjustment by authors	16°
15. Você sente pigarro na garganta? (Do you feel phlegm in the throat?)	0.84	Change to “Sente pigarro na garganta?” (“Do you feel phlegm in the throat?”)—a syntactic adjustment was made by the authors in the original Portuguese, but there is no change in the English version of the question.	17°

(Continued)

TABLE 2.
(Continued)

Question	Content Validity by Item Index (CVI-I)	Decision of Authors	Order of Question in Instrument After Adjustments
16. Você sente que tosse mais hoje em dia do que antes? (Do you feel that you cough more today than in the past?)	0.84	Change to "Sente necessidade de tossir?" ("Do you feel the need to cough?")—adjustment by authors	19°
17. Você sente algo estranho na sua garganta? (Do you feel something odd in your throat?)	0.84	Change to "Você sente algo parado na sua garganta?" ("Do you feel that something is stuck in your throat?")—suggestion of experts	18°
18. Você sente que perde a voz ao longo do dia? (Do you feel that you lose your voice over the course of the day?)	0.78	Change to "Fica sem voz ao longo do dia?" ("Do you lose your voice over the course of the day?")—adjustment by authors	10°
19. Você sente que o som da sua voz muda ao longo do dia? (Do you feel that the sound of your voice changes over the course of the day?)	0.78	Change to "O som da sua voz muda ao longo do dia?" ("Does the sound of your voice change over the course of the day?")—adjustment by authors	11°
20. Você sente que sua voz falha durante as conversas? (Do you feel your voice breaks during conversations?)	0.94	Change to "Tem falhas na voz ao falar?" ("Does your voice break when you speak?")—adjustment by authors	9°

on the validity evidence based on the response processes, and their characteristics are described in the following section.

Validity evidence based on response processes

In this step, the authors attempted to resolve possible misunderstandings in the development of the questions and make relevant adjustments according to the performance and responses of a sample from the target population. A total of 40 subjects aged ≥ 60 years, of both sexes and with clearly preserved cognitive functions, were recruited by convenience sampling from groups of activities for older adults. The sample was stratified in four groups based in demographic and socioeconomic variables according to Table 1. The older adults lived in the cities of Natal and Parnamirim, located in the state of Rio Grande do Norte (RN) in the northeast of Brazil. Natal is the capital of RN and Parnamirim is a neighboring municipality, the third most populous in the state. The distribution of the older adults by gender was defined according to the proportions of male (39 679 or 40.10%) and female (59 253 or 59.90%) older adult residents in both cities, as reported in the 2010 Census³¹ by the "Instituto Brasileiro de Geografia e Estatística" (IBGE).

As there is no standard recommendation in the literature for the allocation of subjects at this stage, the criteria and the number of individuals were defined by consensus among the authors.

Interviews were used to obtain this source of evidence, as this is the most often method mentioned in studies with the same objective as the present study.³² To control for possible bias, all interviews were conducted by the same investigator. After each question was answered by the subject in similar conditions as the future application of the instrument, the interviewer asked about the clarity of the question ("Did you understand the question?") and asked each older adult to repeat the question as he/she understood it ("Please, repeat the question as you understood it"). When the answer had some different element from the original question, the interviewer transcribed the response at the time of interview. The aim of this paraphrase strategy was to identify elements of the development of the questions that could be problematic in meaning or interpretation. This strategy allowed us to analyze individual interpretations, summarize similar interpretations, compare interpretations between respondents and between groups, and analyze the general performance of the questions.

As recommended in the literature,³² indirect strategies were also used. In this case, the interviewer noted, for each question, the response time (time interval between the start of the interviewer's question until the end of all answers related to the question) and any nonverbal responses, such as facial and bodily expressions (ie, discomfort, doubt, impatience, anxiety, or tranquility). The respondents' facial expressions and bodily reactions to each question were observed and registered without scores by the interviewer just after the question was asked. Finally, as previously mentioned, the interviewer asked the respondent to recommend adjustments to questions and suggest other signs or symptoms of voice disorders not covered in the questionnaire but still considered relevant.

All transcripts and reactions registered by the interviewer were further analyzed by the authors of the questionnaire, including the interviewer himself. The analysis was performed on individual question and considered the individual interpretations of the respondents, differences in interpretation between the four groups, reactions perceived by the researcher, absolute number and percentage of older adult respondents that semantically understood the question, and older adult respondents' suggestions. The analysis involved making decisions to delete or change each question, which was done by obtaining the consensus of the four researchers who generated the questions. Once this phase was completed, the third version of the questionnaire was considered complete.

We note that the RAVI were developed in Portuguese language. The English language translations were done by the authors of the article only for publication purposes. There is still no translation and cultural adaptation of the RAVI for the English language.

The present study meets all recommended ethical guidelines for conducting research with human subjects and was approved by the Ethics and Research Committee of the Institution under number 06393312.8.0000.5292.

RESULTS

Validity evidence based on test content

The construct “vocal disorders” was defined by the authors in accordance with the most commonly reproduced concept in literature relating to the voices of older adults¹¹: “vocal disorders refer to when, at any time, the voice fails or the individual perceives his/her voice as not being normal and it interferes with communication.” The initial version derived from the researchers' theoretical and empirical bases consisted of 20 questions. On a more concrete domain, the questionnaire content referred to the sensations and perceptions related to the voice. This first version of the questionnaire was examined by the committee of experts, and the results are summarized in Table 2.

Question 01 had a CVI-I below the recommended level. The experts mentioned flaws in its syntactic construction but considering its theoretical relevance; therefore, the authors decided to keep the question but to modify it as per the experts' suggestions. Question 08 was excluded because it was classified as “inadequate” by >50% of the experts, who attributed little predictive value to this question in relation to the outcome and claimed that the word “intonation” would not be understood by older adults.

Questions 06 and 07 were criticized for representing two extremes of a single issue, which would potentially cause confusion in older adults. By consensus, the instrument authors eliminated question 07 and removed the words “sente que fala ...” (“do you feel that you speak ...”) from question 06; question 06 was then separated into two questions. Question 10 was also split into two questions as described for question 06 and 07 because it represented two extremes of a single issue.

Despite its high CVI-I, question 11 was deleted because some experts pointed out that for many older adults, singing is not a habitual activity; thus, the authors concluded that the

question would not be fully applicable to the target population. In question 17, the phrase “algo estranho na sua garganta” (“something odd in your throat”) was changed to “algo parado na sua garganta” (“something stuck in your throat”). The term “severidade” (“severity”) has various different meanings in Brazilian Portuguese and it was replaced with “gravidade” (in English language it means “severity” too).

In addition to the previously mentioned changes, the questions were reordered during the authors' review of the instrument (last column of Table 2), as per the suggestions of the expert committee. At this point, the questions were sorted into questions about perceptions (questions 01–11) and those about sensations (questions 12–19). Syntactic readjustment—namely, preserving the semantic aspects of the questions—was also performed. The questions were considered to adequately represent the domain, according to a CVI value of 0.80.

The older adults who were interviewed were satisfied with the questionnaire content and did not directly suggest any new questions. However, frequent reporting of an “itch” in the throat caught the attention of the authors, who agreed by consensus to include a question about this topic in the instrument.

Validity evidence based on response processes

The 19-question version (according to last column of Table 2) obtained after the previous analysis of validity evidence based on test content was applied to analyze validity evidence based on response processes. Researchers noted that questions 04, 05, 07, and 08 were not easily understood by the target population, especially by groups B, C, and D. The four questions generated reactions of doubt, confusion, and discomfort that resulted in a longer response time and misinterpretations of the content. The authors observed that the respondents interpreted the intensity and frequency of sensations and perceptions in terms of innate vocal traits instead of problems due to aging. Moreover, the interviewees were frequently under the impression that question 05 had already been answered in question 04, and similarly that question 08 had already been answered in question 07.

Because of their relevance from a theoretical point of view, the authors decided not to eliminate the questions 04, 05, 07, and 08 but to modify them to “Sua voz está forte ou fraca?” (“Is your voice strong or weak?”) and “Percebe que sua voz está mais fina ou mais grossa?” (“Have you noticed whether your voice is high pitched or low pitched?”), respectively. The authors opted to integrate each set of binomial questions into a single question because keeping them separate would be confusing and increase the extent of the questionnaire. Furthermore, we concluded that classification of loudness in strong or weak and classification of pitch in high pitched or low pitched were irrelevant from an epidemiologic perspective.

In question 10, the same previously mentioned negative reactions and interpretations occurred, resulting in a longer response time. The respondents interpreted the question content as “remaining silent during the day,” so the authors changed the question to “Does your voice fade throughout the day?” Hesitation was common in all groups for questions 09 and 11, which led the interviewer to believe that the responses were not authentic. Therefore, the authors combined the two questions

into a single one: “Sua voz piora ao longo do dia?” (“Does your voice get worse throughout the day?”).

Despite being semantically acceptable, questions 12–16 were reexamined by the authors because of the possible confusion between voice and speech. To make it explicit that the content of the questions of the instrument is related to the voice, the words “speak” or “speaking” were excluded from these five questions and the questions were reworded. Question 18 was modified, as some older adults from groups A, C, and D interpreted it differently from the original intent of the question, resulting in greater response times and reactions of doubt. During the interview, it was found that question 19 could be associated with an event occurring naturally or because of other health conditions; thus, it was considered irrelevant to the construct and deleted.

The aspects of time, frequency, intensity, and severity were not properly understood by the target population, resulting in possible not authentic responses and reactions of hesitation and discomfort, as well as an increased total instrument administration time. By consensus, the authors decided to keep only the frequency response scale, which was adjusted to have three response choices: “never,” “sometimes,” and “always.” The third version of the RAVI was completed with 16 questions (Appendix I).

DISCUSSION

The results presented in this study is the first investigating the validity evidence on the basis of the RAVI. Different criteria for measuring the psychometric properties of specific instruments in the field of health have been proposed by different authors and organizations^{33–35}; however, the guidelines recommended by the “Standards”²¹ are considered the most robust and traditional from a psychometric perspective and therefore were adopted by the authors of this article.

In voice research, most instrument validation studies have investigated the psychometric properties of self-assessment questionnaires^{16–20}; however, a systematic review showed that the development methods of some of these instruments were flawed. Thus, their poor psychometric properties were partly due to initial problems in the development of the content and verification of its quality, especially during question or item generation and reduction.²³

The results of the verification of the validity evidence based on content were essential for adjusting the semantic, syntactic, and contextual aspects of the initial version of the RAVI. This step is considered an important source of validity evidence^{36,37} and helps improve questions in terms of both their phrasing and their representativeness and relevance.³⁷

To understand the semantic, syntactic, and contextual aspects of the RAVI, a committee of experts in the field was invited to evaluate the measure. The committee members’ opinions were analyzed both quantitatively and qualitatively, as per the recommendations of previous research,³⁸ and the questions were modified by integrating their suggestions and further empirical analysis by the researchers.³⁷

Despite the large number of questions with acceptable CVI-I according to the experts’ assessments, all the questions had to

be revised, at the very least with minor changes to syntactic construction. This revision was only possible because of the experts’ comments, which encouraged the authors to reflect and rediscuss the details of each question and then make modifications according to what was suggested. The fact that all the questions needed revision highlights the importance of obtaining accurate psychometric properties from a new questionnaire. This is especially important because most existing protocols in voice research have some deficits in the development phase.²³

Branski et al²³ mentioned that it is important to consider the perspective of the target population in developing the questions of a new instrument, in addition to considering the expertise of the researchers and the empirical basis of the literature. This perspective can be obtained through strategies such as focus groups. However, we chose not to use this strategy in developing the questions because of certain peculiarities of the older adult population, described as follows: (1) according to our experience, differentiation between “voice” and “speech” is very inaccurate among Brazilian general population; indeed, this was found when questions possessing the word “speak” or “speaking” had to be modified during the investigation of the validity evidence based on response processes; (2) past studies^{13,14} have pointed out that older adults rarely report voice-related symptoms spontaneously or seek health professionals for this reason, especially because they are resilient in relation to changes in its own voice, accept it as a natural part of aging and do not realize the voice changes as a health problem³⁹; (3) depressive symptoms have proven to be a common and limiting factor for functional activities in older Brazilians,⁴⁰ and the possible accentuated lability and instability of depressed older adults could have influenced the reliability of focus group responses; (4) it is also known that older adults have many comorbidities, so other more disabling health conditions could be more appreciated by them at that time than a voice disorder; (5) Finally, according to the authors’ experience, many older adults are unaware of the relationship between certain symptoms and voice disorders, for example, to feel that the voice is tired.

Therefore, we believe that relevant information would not have been obtained if we do not have an initial version of the questionnaire. We note that the initial version of the questions was not prepared without criteria. It was devised by researchers with extensive expertise and training in the fields of aging and voice, using their considerable experience in combination with their knowledge of gerontology and the support of scientific literature. As described in the study methods and results, the perspective of the older adults was not neglected because they also had the opportunity to contribute via suggestions for adjustments or other relevant signs and symptoms of voice disorders.

Verification of the response process validity proved to be highly relevant for making appropriate adjustments to the RAVI and allowed for modifications based on verbal and nonverbal responses of the target population. Response processes explicitly emerged as a source of validity evidence only in the 1999 edition of the “Standards”, which may explain the small number of validation studies that consider this type of

validity evidence and the absence of guidelines for obtaining it.³²

Validation studies in health care, compared with in other fields, have most frequently explored response processes, and there is notable complementarity between this aspect of validity evidence and that of content,³² as was seen in the RAVI development. The version of the instrument modified according to the experts' opinions required further adjustments following the behavioral and verbal observations obtained from interviews with older adults. Interviews are the most common method for verifying the validity evidence based on the response processes because they best meet the recommendations of the "Standards",²¹ are easy to administer and do not require great financial resources.³² In addition, interviews are recommended for older adults because of the possible difficulties they experience in reading and answering questionnaires on their own.¹

Participant selection was crucial for obtaining accurate results at this stage of the validation process of the RAVI. In this phase, researchers usually consider the characteristics of the target population that are relevant to the study purpose.³² The fact that the RAVI has an epidemiologic perspective was crucial for the authors' decision to create groups by demographic and socioeconomic characteristics. The fact that the groups were composed of different strata of the community enabled us to identify the need to adjust the syntactic, semantic, and contextual characteristics of the questions, thereby making the questions accessible to any respondent.

Other sources of validity evidence and reliability are necessary to continue the validation process of the RAVI. These aspects will be presented and discussed in part II of this study. It should be noted that the version of RAVI presented herein is not final, as further adjustments can be made in subsequent phases. Once the RAVI validation process has been completed, it can be applied on a large scale to facilitate the work of health professionals in mapping the prevalence or incidence of voice disorders among older adults. Furthermore, it can assist in the decision-making of administrators in relation to public policies aimed at preserving human communication among older adults.

CONCLUSIONS

Obtaining validity evidence based on the content and response processes was crucial for the syntactic, semantic, and contextual adjustments needed to adapt the questions of the RAVI to the instrument's ultimate purpose. The results presented herein emphasize the importance of more detailed descriptions of these steps in studies seeking the validity evidence of an instrument. The two aspects of validity evidence studied were complementary and allowed for empirical and theoretical exploration of the various aspects of the questionnaire's development, thereby contributing to the questionnaire's robustness. The version of the RAVI developed in this study will undergo verification by other sources of validity evidence and analysis of reliability to obtain its final version.

REFERENCES

- Kim JW, Nam CM, Kim YW, Kim HH. The development of the Geriatric Index of Communicative Ability (GICA) for measuring communicative competence of elderly: a pilot study. *Speech Comm*. 2014;56:63–69.
- Byles J. The epidemiology of communication and swallowing disorders. *Adv Speech Lang Pathol*. 2005;7:1–7.
- Dehqan A, Scherer RC, Dashti G, Ansari—Moghaddam A, Fanaie S. The effects of aging on acoustic parameters of voice. *Folia Phoniatr Logop*. 2013;64:265–270.
- Etter NM, Stemple JC, Howell DM. Defining the lived experience of older adults with voice disorders. *J Voice*. 2013;27:61–67.
- Merril RM, Roy N, Lowe J. Voice-related symptoms and their effects on quality of life. *Ann Otol Rhinol Laryngol*. 2013;122:404–411.
- Pontes P, Brasolotto A, Behlau M. Glottic characteristics and voice complaint in the elderly. *J Voice*. 2005;19:84–94.
- Pontes P, Yamasaki R, Behlau M. Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop*. 2006;58:151–158.
- Gregory ND, Chandran S, Lurie D, Sataloff RT. Voice disorders in the elderly. *J Voice*. 2012;26:254–258.
- Verdonck-de Leeuw IM, Mahieu HF. Vocal aging and the impact on daily life: a longitudinal study. *J Voice*. 2004;18:193–202.
- Kendal K. Presbyphonia: a review. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:137–140.
- Roy N, Stemple J, Merrill RM, Thomas L. Epidemiology of voice disorders in the elderly: preliminary findings. *Laryngoscope*. 2007;117:628–633.
- Plank C, Schneider S, Eysholdt U, Schutzenberger A, Rosanowski F. Voice and health-related quality of life in the elderly. *J Voice*. 2011;25:265–268.
- Turley R, Cohen SM. Impact of voice and swallowing problems in the elderly. *Otolaryngol Head Neck Surg*. 2009;139:33–36.
- Monini S, Fillippi C, Baldini R, Barbara M. Perceived disability hearing and voice changes in the elderly. *Geriatr Gerontol Int*. 2015;15:147–155.
- Pernambuco LA, Espelt A, Balata PM, Lima KC. Prevalence of voice disorders in the elderly: a systematic review of population-based studies. *Eur Arch Otorhinolaryngol*. Epub 2014 August 23. <http://dx.doi.org/10.1007/s00405-014-3252-7>.
- Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice*. 1999;13:557–569.
- Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI): development and validation. *Am J Speech Lang Pathol*. 1997;6:66–70.
- Deary IJ, Wilson JA, Carding PN, MacKenzie K. VoiSS: a patient-derived Voice Symptom Scale. *J Psychosom Res*. 2003;54:483–489.
- Ma EP, Yiu EM. Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res*. 2001;44:511–524.
- Bach KK, Belafsky PC, Wasylik K, Postma GN, Koufman JA. Validity and reliability of the glottal function index. *Arch Otolaryngol Head Neck Surg*. 2005;131:961–964.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. New York: American Educational Research Association; 2014.
- Carretero-Dios H, Perez C. Standards for the development and review of instrumental studies: considerations about test selection in psychological research. *Int J Clin Health Psychol*. 2007;7:863–882.
- Branski RC, Cukier-Blaj S, Pusic A, et al. Measuring quality of life in dysphonic patients: a systematic review of content development in patient-reported outcomes measures. *J Voice*. 2010;24:193–198.
- Alexandre NM, Coluci MZ. Content validity in the development and adaptation processes of measurement instruments. *Ciênc Saúde Coletiva*. 2011;16:3061–3068. Portuguese.
- Sireci SG. The construct of content validity. *Soc Indicators Res*. 1998;45:83–117.
- Sireci S, Faulkner-Bond M. Validity evidence based on test content. *Psicothema*. 2014;26:100–107.
- Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs*. 2006;53:205–212.
- Gallardo RY, Olmos RC. The Delphi method and the investigation in health services. *Ciênc Enferm*. 2008;14:9–15 [Spanish].

29. Falzarano M, Pinto GP. Seeking consensus through the use of the Delphi technique in health sciences research. *J Allied Health*. 2013;42:99–105.
30. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29:489–497.
31. Instituto Brasileiro de Geografia e Estatística [internet]. Censo 2010. 2010. Available at: <http://censo2010.ibge.gov.br/>. Accessed March 20, 2013.
32. Padilla JL, Benítez I. Validity evidence based on response processes. *Psico-thema*. 2014;26:136–144.
33. Scientific Advisory Committee of Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res*. 2002;11:193–205.
34. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737–745.
35. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurements properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34–42.
36. Squires JE, Estabrooks CA, Newburn-Cook CV, Gieri M. Validation of the conceptual research utilization scale: an application of the standards for educational and psychological testing in healthcare. *BMC Health Serv Res*. 2011;11:107.
37. Delgado-Rico E, Carretero-Dios H, Ruch W. Content validity evidences in test development: an applied perspective. *Int J Clin Health Psychol*. 2012;12:449–460.
38. Magasi S, Ryan G, Revicki D, et al. Content validity of patient-reported outcome measures: perspectives from a PROMIS meeting. *Qual Life Res*. 2012;21:739–746.
39. Souza IA, Massi G. The speech therapy health based on the institutionalized elderly speech. *Rev CEFAC*. 2015;17:300–307.
40. Lima AL, Lima KC. Activity limitation in the elderly people and inequalities in Brazil. *Open Access Libr J*. 2014;1:e739.

APPENDIX I

Rastreamento de Alterações Vocais em Idosos (RAVI; “Screening for Voice Disorders in Older Adults”; preliminary version obtained from validity evidence based on test content and response processes)*

Questões (Questions)	Sim (Yes)	
	Não (No; 0)	<div>Às Vezes (Sometimes; 1)</div> <div>Sempre (Always; 2)</div>
(1) Sua voz lhe incomoda? (Does your voice bother you?)		
(2) As pessoas se incomodam com a sua voz? (Does your voice bother other people?)		
(3) Sua voz é rouca? (Is your voice hoarse?)		
(4) Sua voz está forte ou fraca? (Is your voice strong or weak?)		
(5) Sua voz some ao longo do dia? (Does your voice fade throughout the day?)		
(6) Sua voz piora ao longo do dia? (Does your voice get worse throughout the day?)		
(7) Percebe tremor na sua voz? (Do you feel that your voice is shaky?)		
(8) Percebe que sua voz está mais fina ou mais grossa? (Have you noticed that your voice is high pitched or low pitched?)		
(9) Sente que faz esforço para a voz sair? (Do you have to strain to produce your voice?)		
(10) Sente cansaço na voz? (Do you feel that your voice is tired?)		
(11) Sente sua garganta seca? (Do you feel that your throat is dry?)		
(12) Sente coceira na garganta? (Do you feel your throat itch?)		
(13) Sente queimação, ardência na garganta? (Do you feel burning or irritation in your throat?)		
(14) Sente pigarro na garganta? (Do you feel phlegm in the throat?)		
(15) Sente dor na garganta? (Do you feel pain in your throat?)		
(16) Sente alguma coisa presa na garganta? (Do you feel as if something is stuck in your throat?)		

* The English language translations were done by the authors of the article only for publication purposes. There are still no translation and cultural adaptation of the RAVI for the English language.